Vol. 2, Issue I, June 2021 **ISSN: 2582-3310**

Development of *in-silico* based Intron Length Polymorphism (ILP) Marker in medicinally important *Catharanthus roseus plant*

Shivangi Bansal and Raghvendra Kumar Mishra*

Amity Institute of Biotechnology, Amity University Madhya Pradesh Gwalior, India

*Corresponding Author: rkmishra@gwa.amity.edu

Abstract

Catharanthus roseus (Madagascar periwinkle) is a well-known plant with high medicinal value. Traditional methods of crops or plants improvement like conventional breeding (classical breeding) program have been largely limited by self-compatibility and heterozygosity. Recently, DNA based molecular marker-assisted breeding has increased the speed of breeding program; reduce the manpower as well time to develop elite varieties. Although the development and application of large-scale markers has been reported in *Catharanthus roseus*, but till now, Intron-length polymorphim markers (ILP) was not reported. For the development of intron length polymorphism (ILP) markers, 22,867 EST sequences were retrieved from NCBI database and pre-processed. The overlap sequences were identified and only single and coatings sequences were used in the study. ILP markers were designed by comparing the EST sequences with the available genomes of model plants. Approximately 38 primer pairs were designed for the *C. roseus* from flanked potential intron positions. The BLASTx (EST) analysis of 22,867 express sequence tags suggested that the 55.5% function for ESTs sequence edit them into 2 different functional categories. The developed ILP primers representing the different genic regions of *C. roseus* and will be able to identify polymorphic characters and gene position, diversity analysis and study of transferability.

Keywords- *Catharanthus Roseus*, Intron-Length Polymorphism (ILP), molecular marker, genome, ESTs, PIP database

Vol. 2, Issue I, June 2021 **ISSN: 2582-3310**

INTRODUCTION

The *Catharanthus roseus* is an important medicinal plant in the Apocynaceae family and it is widely used as a source of chemotherapeutic drugs. This family mainly contains herbs and small shrubs. It has smooth marginal leaves, the flowers are found in leaf axils, which are born separately or in pairs on very short stems, and it has another distinctive feature which is the potent milky sap (1).

Catharanthus roseus grows to a height of 20-60 cm, the flowers of this plants are pink, white or rosy-purple. The flowers have a base tube 2.5-3.0 cm long, about 2.0-5.0 cm in diameter, with five petal- like lobed. There are 5 sepals, 2-6 mm long, narrow, usually with pubescent.

Fig1- Picture of *Catharanthus roseus*

Flowering is usually happening during the summer months (March to May months). Very high temperature is not suitable for flowering. Leaves are simply oppositely arranged on the stem, with the entire margin if the leaves, also the plants have an imperceptible and indistinguishable fruit. These Plants are either propagated by seeds or by vegetative methods. Room temperature (25^0C), dark conditions and low water are the suitable conditions for the germination of the seeds (2,3).

The flowers of *Catharanthus roseus* are pollinated by butterflies and moths, this species is self-compatible, although under normal conditions self-pollination is rare. This plant also can be use as ornamental plant in garden and homes across warmer places, or can be grow in glasshouse throughout cold season. *Catharanthus roseus*, better known as the Periwinkle of Madagascar, is native to the island of Madagascar in the Indian Ocean. Madagascar is located on the east coast of South Africa. The Periwinkle is a perennial plant that is very common in tropical and subtropical forests (2, 4).

Vol. 2, Issue I, June 2021 **ISSN: 2582-3310**

Till date several medicinal plants were already identified around the globe. However it was reported that *C. roseus* is one of the most important medicinal plants due to avainlability of more than 200 secondary metabolites. Every part of plant (stem, root, leaf) are highly useful. They are the rich source of Alkaloids (TIAs). In addition to alkaloids, *Catharanthus roseus* produces a wide range of phenolic compounds, including C6C1 compounds such as 2,3-dihydroxybenzoic acid, as well as phenylpropanoids, such as cinnamic acid derivatives, flavonoids. The formation of these compounds in *C. Roseus* is reconsidered, as well as their biosynthesis and regulation of the path. Both types of compounds compete with the biosynthesis of indole alkaloids (4-6).

Introns are very important building blocks of any genomes and scattered throughout the genome. These are the non-coding sequences present in the gene that are transcribed, but removed during the pre-processing. For example, introns make up ~25% of the human genes, respectively (7). In general, introns have little functional significance, although some insertions can affect levels of gene expression. Consequently, introns are more variable than coding sequences.

In plant, fluctuations or polymorphisms can be identified by genetic markers (DNA based marker) with the help of PCR based technologies, which are very useful tools for genetic research (for example, building genetic maps, defining genes or locations for quantitative properties). With the help of DNA markers and mapping population, several genetic maps were already developed in several crops (8, 9-12). On the basis of sequencing technologies and development of new tools, several DNA based markers have been developed, such as microsatellite or simple sequence repeat (SSR), Insilco based markers, single nucleotide polymorphism (SNP) and ILP marker etc. (12-18).

Variation in Intron sequences can also be used to detect the polymorphism. They have been used successfully in mapping research genetics and population genes. It can be easily detected by PCR. To amplify introns by PCR, primers can be designed from flanking exons. This approach is called exon-primed intron-crossing PCR (EPIC-PCR). This approach provided a new method to identify and amplify the DNA sequences. It was reported previously that exon sequences are highly conservative; due to this unique character primers designed from these flanking sequences will be highly use full in several studies.

Because of their unique properties, ILP markers are unique because they are gene specific, co-dominant, hyper-variable, neutral, convenient and reliable. In addition to the sequence tagged sites markers, ILP markers also have transferability ability to amplify adjacent plant species.

To facilitate the direct advancement of ILP markers (13, 17) developed an online database called PIP (Potential Intron Polymorphism) to provide detailed information on various types of markers, indicators and homologous relationships. Despite these advantages, no reports are available till

Vol. 2, Issue I, June 2021 **ISSN: 2582-3310**

date on the uses of ILP markers in medicinal important *C. roseus* plant. The primary objective of the present study was to development of ILP from the publicly available *Catharanthus* ESTs and their Characterization.

Materials and Methods

Molecular markers are most popularly used for estimation of polymorphisms, relatedness & mating system parameters, genotype characterisation in medicinal plants.

Hence, ILP was developed for *Catharanthus roseus*. Therefore, the ILP markers were developed for the same.

Identification of EST sequences

First, the *Catharanthus roseus* EST sequences were retrieved in FASTA format from the NCBI (https://www.ncbi.nlm.nih.gov/) i.e. National Centre for Biotechnology Information advanced science & health by providing access to biomedical & genomic information.

- Open the web browser & enter NCBI in the query, & click on the first link for the NCBI.
- Enter *Catharanthus roseus* in the query box of the page & select Nucleotide from the database, further click on the search button.
- Click on EST from the search result, to select ESTs.
- Go down on the page & click on Send file & select File, then select FASTA in the format & accession in sort by, & click on create file.
- Hence the EST file of the desired EST is developed and downloaded.

Pre-processing of the sequences

The pre-processing was done for the FASTA sequences retrieved from the web, with the help of the web server of the software named as CAP3 (http://doua.prabi.fr/software/cap3) to identify the unique EST sequences.

The Cap3 algorithm computes overlaps between sequences & then joins the reads in decreasing order of overlaps to form Contigs.

- Opened http://doua.prabi.fr/software/cap3 & entered the sequences from the FASTA file retrieved from NCBI Database.
- "Cap3 gave the two files after the pre-processing, that is Contigs & the Single tone sequences & therefore, the further processes were done separately for these."

Selection of candidate EST sequences

Vol. 2, Issue I, June 2021 **ISSN: 2582-3310**

Basic Local Alignment Search Tool (BLAST; https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to identify the specific functions of Non-redundant ILP containing EST sequences. On the basis of blast hit, the homologous genomic regions were identified. BLAST searches were performed to provide complete coverage across the genome sequence.

- Open https://blast.ncbi.nlm.nih.gov/Blast.cgi, & enter the sequences from the file.
- Submit the sequences, by setting the desired parameters.
- Among the various hits, selected the one with the most query coverage, identity & highest scores.

Data Analysis

The data collected from the CAP3 & BLAST were than analysed fir the identification and the characterization of ILP.

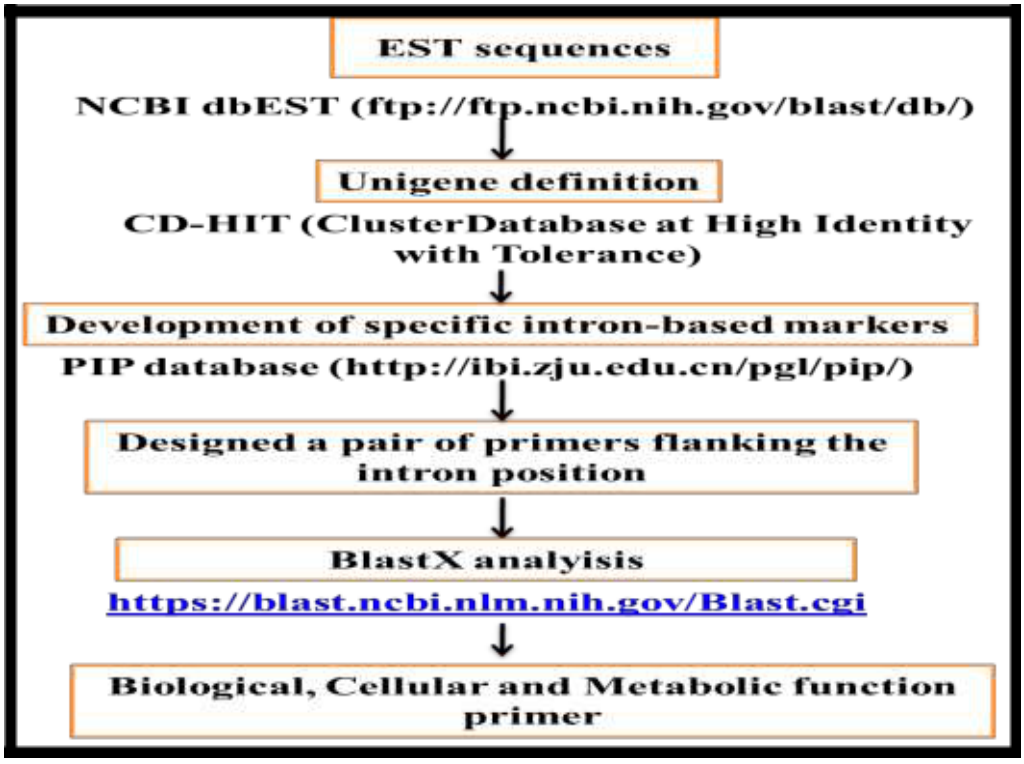


Fig 2- Flow Chart of Methodology

Vol. 2, Issue I, June 2021 ISSN: 2582-3310

Results and Discussion

EST assembly

We downloaded a set of 22,867 ESTs from *Catharanthus roseus* EST-database available at NCBI: https://www. ncbi.nlm.nih.gov/.

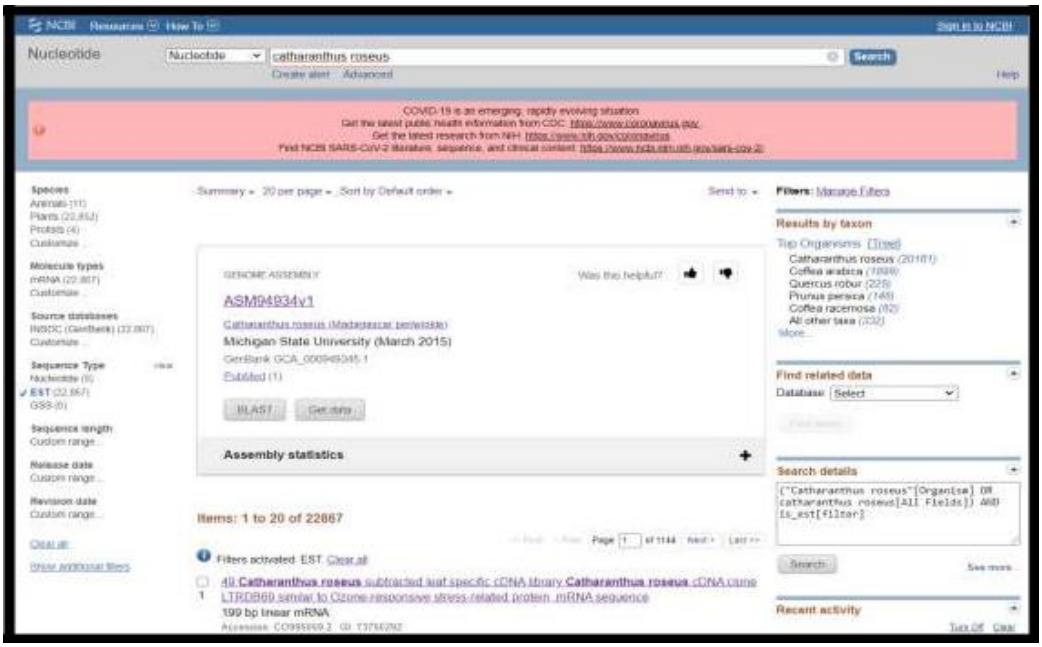
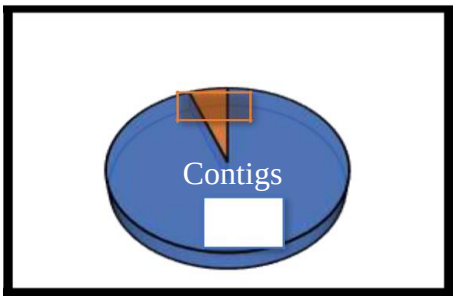


Fig 3: EST Sequences of *Catharanthus roseus* retrieving from NCBI Database (17-07-20)

Pre-processing of sequences

EST sequences of *Catharanthus roseus* were downloaded from ftp://ftp.ncbi. nih.gov/blast/db/. Approximately 22,867 sequences were pre-processed to remove the overlapping sequences. Around 18,992 singletone and 1127 contings were identified and selected by CAP3 software (12, 19). Further these sequences were used to identify SSR containing sequences.



Singletone

Fig 4 - EST sequences of *C.roseus*

Vol. 2, Issue I, June 2021 **ISSN: 2582-3310**

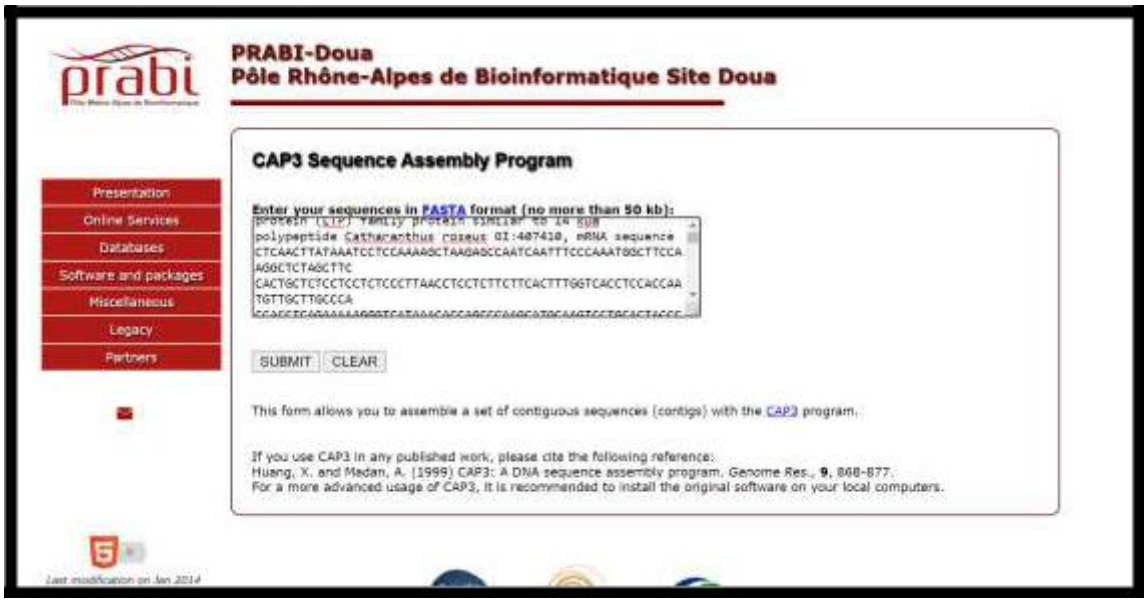


Fig 5- Prabi CAP3 (17-07-20)

ILP mining and primer designing

Unique sequences were processed for developing ILP primers flanking introns using *Catharanthus roseus* singleton genomic sequences. PIP identifies exon intron boundaries and predicts suitable primers flanking intronic regions (Table 1).

Characterization of Primers

Developed ILP primer pairs were characterized by using BLASTX searched and analysed. A cut-off bit score of GC % content above 50% and an E-value of 1e-05 were considered optimum for BLASTX analysis.

The supposed functions of the ILP markers are assigned to perform the BLASTX search for the corresponding markers that contain the EST strings in the NCBI database(http://blast.ncbi.nlm.nih.gov/Blast.cgi) with standard search parameters.

Vol. 2, Issue I, June 2021 **ISSN: 2582-3310**

Table 1 List of primers and their amplification characteristics

S.No.	ID	Tm (˚C)	GC%	Forward primer	Reverse Primer
1	AT4G37930	58	39	GACAACCACTTAGTTTTGGTGAA	CCAGGAACAGTGTTTTTGTTAGC
2	AT1G23490	60	55	TCTCCAGAAGTGGCACAGGT	TCTTCCCAATGCAATGAATG
3	AT5G12250	57	33	ATTTGCTTCAGGACTCTTAAACTT	AATTGAGCTGACCAGGGAAC
4	AT5G54680	60	43	GAGGATTCATGTTCTTCAGTTGC	AAAGAACTTAAAGCTGAGAAAAAC
5	AT2G36580	58	43	GCAATTAAAGTCAAGGCATCC	AAGACGAGGGATGACAACAGA
6	AT5G17770	58	39	TCCTTCCTTGAGGATACATCTTT	AGTCAAGGTGAGGAGGTGATT
7	AT5G17770	57	47	AATCACCTCCTCACCTTGACT	TCAGCCATAATGTAGCAAAGTTTA
8	AT4G39910	60	34	GGGTGGCTTTTTAATTTTCATTC	GAAGAACTTTAGTTCCACCGAGA
9	AT1G12240	59	40	TCCTTGGCATGGAATTTTTC	GACAAGAAGGAATTCGCTCAA
10	AT5G25150	62	63	CGGCCGAGGTACAGCACTA	AGCCACCAGCAACAAGAGAT
11	AT1G13950	61	56	CGGCACCATCCGTAAGAAC	ACTTGGCATGACCGTGCT
12	AT2G39390	59	45	CTGACTGTAATTTCGCAGAAGC	GGATGGCTCTGGTGTTCTTC
13	AT1G51160	59	50	ATGAACACCCTTGTCCAGGT	AAGCTCATGTTTGGCTTGCT
14	AT5G16960	59	35	TAAATATCAATGCCCTCTGGAAA	TGAAGAACAAATTTGGCTTTGAT
15	AT1G76630	59	55	GCCGAGGTACCTTCTCTGAA	TGGAGTGTGAAGGCAATTAGG
16	AT2G19790	59	50	ATGATGTCCAGTTCGCACAC	TTGGTGGGAGTTGACAATGA
17	AT2G19790	59	45	TTCATTGTCAACTCCCACCA	CTCGCACCGAACAACAGTG
18	AT4G26900	59	39	TGCACCTGATCTGAAATACTTTG	TTGTACCGTTGACAGTTGGTG
19	AT4G26900	59	61	CCCCCACCAACTGTCAAC	TTACGGGTTTTCGAGACTTCC
20	AT4G26900	60	60	GTCCCCAAGAGGGAAGTCTC	AACCTTGGCAAGCCAGTAGA
21	AT4G35450	60	57	ACTCCTTGCGTCCGTAACC	GGAGAGGTGAAATGTGCTCAG

Vol. 2, Issue I, June 2021 **ISSN: 2582-3310**

Table 2 Characteristics of EST-derived ILPs for *Catharanthus roseus*.

Primer name	Gene Bank accession no.	Expected Product Size (bp)	Query length	E-value	Putative function
AT4G37930	CO995058.1	110	400	2e-56	hypothetical protein EE612_020327 [Oryza sativa]
AT1G23490	DT527671.1	105	88	3e-45	hypothetical protein SETIT_3G198500v2 [Setaria italica]
AT5G25150	EG562736.1	110	204	7e-24	hypotheticalproteinC5167_049086[Papaver somniferum]
AT1G13950	EG562668.1	101	432	9e-78	eukaryotic translation initiation factor 5A-4 [Cannabis sativa]
AT1G76630	EG562602.1	107	384	3e-36	TPRrepeat-containingprotein[Handroanthus impetiginosus]
AT4G26900	EG562578.1	112	572	6e-105	hypothetical protein CISIN_1g018578mg [Citrus sinensis]
AT1G69620	CX119705.1	100	565	9e-58	Select seq ref|XP_004238799.1| 60S ribosomal protein L34 [Solanum lycopersicum]
AT1G62040	EG562485.1	113	505	6e-40	putative microtubule-associated protein [Oryza sativa]
AT5G52660	EG562479.1	100	676	8e-60	protein REVEILLE 6 [Solanum tuberosum]
AT4G17300	EG562465.1	107	657	6e-84	asparagine--tRNA ligase, chloroplastic/mitochondrial [Solanum pennellii]
AT1G76160	EG562441.1	108	446	6e-70	L-ascorbate oxidase homolog [Nicotiana tomentosiformis]
AT5G27850	EG562407.1	109	535	8e-106	putative 60S ribosomal protein L18-1 [Hibiscus syriacus]
AT1G68370	EG562374.1	101	534	1e-116	chaperoneproteindnaJ15-like[Nicotiana tomentosiformis]
AT2G18110	EG562290.1	108	480	2e-29	elongation factor 1-delta 1 [Citrus sinensis]
AT2G36360	EG562271.1	108	421	4e-58	rab9 effector protein with kelch motifs isoform X1 [Helianthus annuus]
AT3G23390	EG562261.1	108	399	3e-51	hypothetical protein F8388_026239 [Cannabis sativa]
AT5G53560	EG562232.1	111	529	1e-43	cytochrome b5 isoform E-like [Solanum tuberosum]
AT4G27960	EG561828.1	111	504	2e-98	putative aminoacyltransferase, E1 ubiquitin-activating enzyme [Lupinus albus]

Volume II, Issue I, June 2021 **ISSN: 2582-3310**

Data mining for development of ILP markers

Around 22,867 ESTs sequences of C. roseus were extract from the EST database, which were collected in 20,119 unique sequences (18,992 Contigs and 1127 singletons) and identified 38 ILP primers. In addition, ILP indicators were extracted from the PIP database. Thus, the PIP database served as a potential resource marker for mining - development of new markers; The remaining strings could not be used to design ILP markers, as they were not able to meet the initial design criteria.

As introns are a non-coding DNA sequence in genes that are not functionally important in plant metabolism, although some genes can affect the level of gene expression. In addition, in the genome they are more variable than the coding sequences due to the total selective pressure in intense regions, much less than in exonic regions (13, 17). Although Huang et al. (16) showed that ILP markers are more versatile than other markers, but there are very few polymorphic ILP markers. In addition, these markers have fewer crosses in wild relatives compared to ILP markers. Thus, these ILP markers are suitable for characterizing wild relatives.

Functional annotation of ILP

BLASTX analyses of the 2,867 EST sequence took on almost many defined functions for ILP markers, and some had nothing in common with previously sequenced genes. The function-based ILP markers were grouped into five main categories with defined function (55.5%) (Figure 2). The largest category (37.2%) contained EST sequences with hypothetical /undeclared/ putative functions. The second largest class (22.6%) included the photosynthesis gene. The stress-related gene (12.5%) ranked third, followed by defence (11.2%), followed by secondary metabolism (8.1%), transformation factors (6.1) and primary metabolism (2, 3%) (Figure 2). *Catharanthus roseus* is a potentially abiotic stress-tolerant plant, particularly for drought salinity, studying EST sequences with hypothetical / undeclared / putative functions (37.2%) and with those that had no resemblance to previously sequenced genes (45, 5%) can introduce new details of stress tolerance mechanisms.

Volume II, Issue I, June 2021 **ISSN: 2582-3310**

Functional Classification

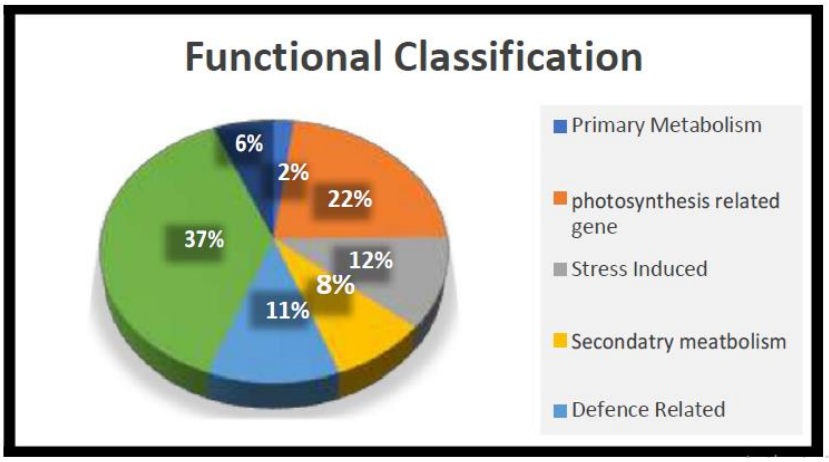
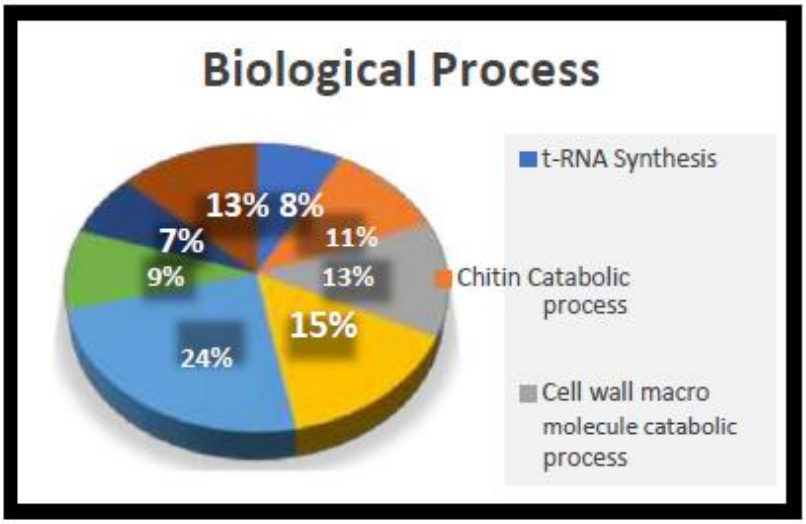


Fig 6- Functional classification of ILP markers containing ESTs/genes. The unique genes were grouped into seven functional groups.

Biological and Molecular Function:

The putative function of the ESTs was attributed to the BLASTx analysis to classify them in two groups based on homology: (a) biological and (b) molecular similar to previous study (12). The significant value of this work is that ILP have shown homology as a target function, demonstrating a good approach for using these ILP sites as a molecular marker to saturate primary and secondary metabolic pathways in plants (13). The polymorphism analysis and transferability of the ILP markers showed the value of the developed indicators. The potency of the ILP developed between species provides a good opportunity to study unknown medicinal plants. The high affinity for these initial pairs, polymorphism and transferability suggests that the markers developed in this study are very useful in auxiliary marker selection, genetic diversity studies, link mapping and comparative analysis. The work presented here complements efforts to develop a well-matured molecular map of *C. roseus*. Recently developed ILPs are informational studies and are a valuable source of gene-based ILP markers.

Volume II, Issue I, June 2021 **ISSN: 2582-3310**



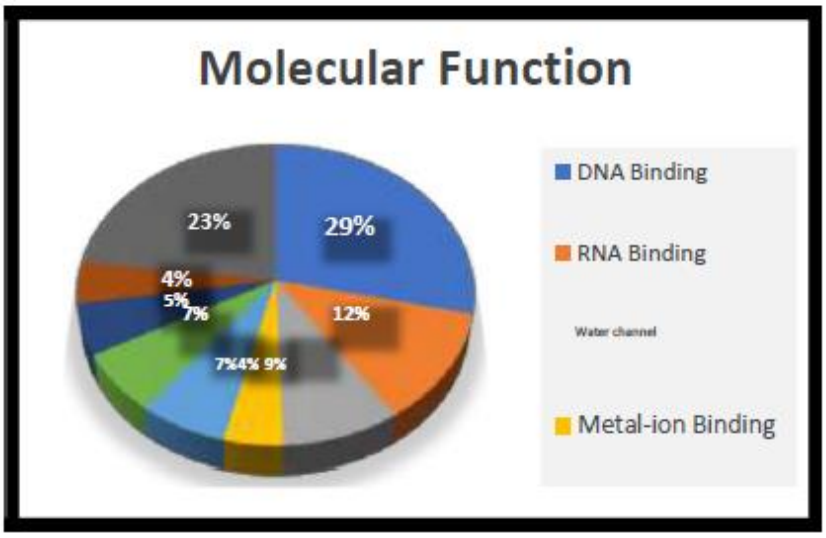


Fig 7. Classification of EST-ILPs on the basis of significant match with putative proteins (A) Biological putative proteins and (B)Molecular Function proteins.

Conclusion

Catharanthus roseus, commonly known as "periwinkle (evergreen)", is a highly studied plant due to its medicinal properties. The secondary metabolites produced by this plant are antileukemic (vincristine, vinblastine) and antihypertensive (ajmalicin

Volume II, Issue I, June 2021 **ISSN: 2582-3310**

and serpentine) (20). However, extremely low yields prevent the widespread use of these alkaloids for therapeutic purposes. Molecular marker-based techniques were used to map high-yielding varieties. Currently, ILPs are evolving as a powerful class of molecular markers due to their availability, hyper variability, high availability analysis, high polymorphism, portability compared to other relevant availability indicators. Currently, ILP markers are used in the selection high-yielding varieties, molecular mapping and analysis of quantitative properties. Apart of these unique properties the development of ILP markers are also required lower cost, less time and highly informative (13,17). Although the development of ILP markers has been reported in some plant species (13,17) but in Catharanthus, this is the first report of development of ILP primers. During this study, 38 primers were developed and characterized.

The ILP primers developed from *C. roseus*, can be used for everyone to characterize the identification of a *Catharanthus* species and can be used to characterize and identification of *C. roseus* genotype, genetic diversity, genetic improvement and molecular DNA deformation test of *Catharanthus* species. In addition, the high level of transmissibility of their cross-breed species will increase our understanding of the penetration of genes, evolutionary relationships, among wild relatives of *Catharanthus* species.

Acknowledgement

We wish to express our sincere acknowledgement to Dr. Ashok Kumar Chauhan, President, RBEF parent organization of Amity University Madhya Pradesh (AUMP), Dr. Aseem Chauhan, Additional President, RBEF and chairman of Amity University Gwalior Campus, Lt. Gen. V.K. Sharma, AVSM (Retd.), Vice Chancellor of AUMP Gwalior Campus, for providing necessary facilities, their valuable support and encouragement throughout the work.

References

Volume II, Issue I, June 2021 **ISSN: 2582-3310**

1. Edward, F.G., & Howe, T. (2013). *Catharanthus roseus* Periwinkle, Madagascar periwinkle. Retrieved January 20, 2014, from http://edis.ifas.ufl.edu/fp112.
2. Askaran, K. K., Kulkarni, R. N., Kumar, S. S. and Sreevalli, Y. Y. (2001). The mechanism and inheritance of intraflower self-pollination in self-pollinating variant strains of periwinkle. Plant Breeding. 120: 247-250.
3. Kulkarni, R. N., Baskaran, K.., & Jhang, T. (2016). Breeding medicinal plant, periwinkle [Catharanthus roseus (L) G. Don]: a review. *Plant Genetic Resources*, 14(04), 283– 302.
4. Mustafa, N.R., & Verpoorte, R,(2007). Phenolic compounds in *Catharanthus roseus*. *Phytochem Rev* 6, 243-258.
5. Shukla, A.K., Shasany, A.K., Gupta, M.M., & Khanuja, S.P.S. (2006). Transcriptome analysis in Catharanthus roseus leaves and roots for comparative terpenoid indole alkaloid profiles. *J of Exp Bot* 57, 3921-3932.
6. Kaushik, S., Tomar, R.S., Gupta, M., Mishra, R.K. (2017). An overview of Catharanthus roseus and medicinal properties of their metabolites against important diseases. *European Academic Research* 5(2), 1237-1247.
7. Berget, S.M., Moore, C., & Sharp, P.A. (1977). Spliced segments at the 50 terminus of adenovirus 2 late mRNA, Proc. Natl Acad. Sci. USA, 74, 3171–3175.
8. Williams, J. G. K., Kubelik, A. R., Livak, K. J., et al. (1990), DNA polymorphism amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18, 6531–6535.
9. Thiel, T., Michalek, W., Varshney, R., & Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.). *Theor Appl Genet* 106(3), 411–422.
10. Gupta, P.K., Rustgi, S., Sharma, S., Singh, R., Kumar, N., & Balyan, H.S. (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol Gen Genom*, 270,315-323.
11. Gupta, S., & Prasad, M. (2009). Development and characterization of genic SSR markers in *Medicago truncatula* and their transferability in leguminous and non-leguminous species. *Genome* 52, 761-771.

Volume II, Issue I, June 2021 **ISSN: 2582-3310**

12. Mishra, R.K., Gangadhar, B.K.H., Yu, J.W., Kim, D.H., Park, S.W. (2011). Development and characterization of EST based SSR markers in Madagascar periwinkle (*Catharanthus roseus*) and their transferability in other medicinal plants. *Plant Omics* 4(3), 154-162.
13. Gupta, S., Kumari, K., Das, J., Lata, C., Puranik, S., & Prasad, M. (2011). Development and utilization of novel intron length polymorphic markers in foxtail millet (*Setaria italica* (L.) P. Beauv.). *Genome*. 54(7), 586-602.
14. Gupta, S.K., & Gopalakrishna, T. (2010). Development of unigenederived SSR markers in cowpea (*Vigna unguiculata*) and their transferability to other Vigna species. *Genome* 53, 508-523.
15. Gong, Y.M., Xu, S.C., Mao, W.H., Hu, Q.Z., Zhang, G.W., Ding, J., & Li, Y.D. (2010) Developing new SSR markers from ESTs of pea (*Pisum sativum* L.). *J. of Zhejiang Univ*. 11,702-707.
16. Huang, M., Xie, F., Chen, N., Zhao, X., Jojee, L. et al. (2010). Comparative analyses of genetic diversity and structure in rice using ILP and SSR markers. *Rice Sci* 17, 257–268.
17. Muthamilarasan, M., Suresh, B.V., Pandey, G., Kumari, K., Parida, S.K., & Prasad, M. (2014). Development of 5123 intron-length polymorphic markers for large-scale genotyping applications in foxtail millet. *DNA Res* 21(1), 41–52.
18. Villano, C., Esposito, S., Carucci, F., Iorizzo, M., Frusciante, L, Carputo, D., & Aversano, R. (2019). High-throughput genotyping in onion reveals structure of genetic diversity and informative SNPs useful for molecular breeding. *Mol Breed* 39(1), 5.
19. Altschul, S., Madden, T., Schaffer, A., et al. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*, 25, 3389–3402.
20. Vander, Heijden, R., Jacobs, D.I., Snoeijer, W., Hallard, D., & Verpoorte, R. (2004). The Catharanthus alkaloids: pharmacognosy and biotechnology. *Curr Med Chem* 11(5), 607-628.

Evolutionary relatedness of human pathogenic bacteria based on conserved and structural gene sequences

Manish Kumar* and Raghvendra Saxena